

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Om nucleic - nucleic search, using sw model

Run on: March 15, 2002, 04:14:19 ; Search time 3106.14 Seconds

(without alignments)
23342.510 Million cell updates/sec

Title: US-09-652-292-1

Perfect score: 4395

Sequence: 1 ggggggttccttgcacggcc.....attatttgtaaaaaaa 4395

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters:

2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_bt:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pt:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_sy:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_rod:*

33: em_htg_hum:*

34: em_htg_inv:*

35: em_htg_rod:*

36: em_htg_other:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4395	100.0	4396	9	AF244053 Homo sapi
2	2465	56.1	4167	9	AF321240 Novel hum
3	2424	55.2	4126	9	AL137188 Novel hum
4	1810	41.2	127418	9	HS28H20 Human DNA
5	182	4.1	385	6	AX150151 Sequence
6	156	3.5	411	11	G31518 Human STS S
c	7	99	2.3	1520	9 BC010033
c	8	99	2.3	1524	9 BC005060
c	9	99	2.3	175599	2 AC009133
c	10	2.3	181705	2 AC023831	
c	11	67	1.5	70875	9 AC023105
c	12	64	1.5	185926	2 AC027632
c	13	63	1.4	108162	2 AC020960
c	14	63	1.4	113982	2 AF192303
c	15	63	1.4	122772	2 AC008874
c	16	63	1.4	164774	2 AC019104
c	17	63	1.4	169241	2 AL45987 Homo sapi
c	18	63	1.4	176232	2 AL354659 Homo sapi
c	19	63	1.4	198027	2 AC008515 Homo sapi
c	20	63	1.4	298734	2 AC008972
c	21	62	1.4	3351	9 AF048726
c	22	62	1.4	75778	9 AL162260 Human DNA
c	23	62	1.4	150009	2 AC022282
c	24	62	1.4	150856	9 AC087072
c	25	62	1.4	154611	2 AC068893 Homo sapi
c	26	62	1.4	166872	2 AC021033
c	27	62	1.4	169725	2 AC073367
c	28	62	1.4	183029	2 AC021030
c	29	62	1.4	186581	2 AC021031
c	30	62	1.4	193877	2 AC009386
c	31	62	1.4	195176	2 AL353576 Homo sapi
c	32	62	1.4	196822	2 AC021041
c	33	62	1.4	196933	2 AC079854
c	34	62	1.4	197480	2 AC0668380
c	35	62	1.4	208561	2 AC069340
c	36	61	1.4	63545	2 AC087639
c	37	61	1.4	161204	2 AC025772
c	38	61	1.4	166211	2 AC079135
c	39	61	1.4	166263	2 AC019047
c	40	61	1.4	177112	2 AC007498
c	41	61	1.4	185727	2 AC090734
c	42	61	1.4	189036	2 AC019142
c	43	61	1.4	207481	2 AL074138
c	44	60	1.4	85173	9 AL449143 Human DNA
c	45	60	1.4	98404	2 AL355676 Homo sapi

ALIGNMENTS

RESULT	1	LOCUS	AF248053	mRNA	PRI
DEFINITION		Homosapiens	glucose transporter (GLUT10)	mRNA	12-APR-2001
ACCESSION		AF248053		complete cds.	
VERSION		AF248053.1			
KEYWORDS					
SOURCE					
ORGANISM					
Human					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;					
Mammalia; Buthidae; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
JOURNAL					
REFERENCE					
AUTHORS					

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

TITLE	Direct Submission
JOURNAL	Submitted (22-MAR-2000) Molecular Genetics, Wake Forest University Medical Center Boulevard, Winston-Salem, NC 27106, USA
ATTRIBUTES	
SOURCE	<p>Location Qualifiers</p> <p>1. 4396</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/chromosome="20"</p> <p>/map="20q12-q13.1"</p> <p>1. 4396</p> <p>/gene="Gutr10"</p> <p>251. 1876</p> <p>/gene="Gutr10"</p> <p>/codon_start=1</p> <p>/product="glucose transporter"</p> <p>/protein_id="AAK31911.1"</p> <p>/db_xref="GI:13603277"</p> <p>/translation="MCHSPPVVPLQASVSVLGLGLPFGYELAVISGALPLQLDGLSLV LEQFLVGSLLLGALLASLGGLFLDCYGRKAQILGSNVLAGSITLGAGSLANLV LGRAVVGFLISLSSAACIYVSEPLPQAGTIVLILSYAINYLAGT" PWNQHMHFMWATAFPAVLQSLSLFLPQACTETATKDIPLQGGEAPLKLGGRPRIFSF LDLFRAIDNMRGRTIVGLGLVFLQOLGPVLYCASTITSSVGRIGGGASVASYLV GAVKVAATIATMGIVDRAASRALLAGAIMALMSVGLYSFAVMDSGSPCLAVPN ATGPGPPUFARGHALLRWPAICLMVEVSFSFGGPVTVLVEYPEVGRATAF CSNFWAANLFIISLUSFLDQNLSTQHNTGIPSRLEASAS</p> <p>DOQFQKRERTLSPECHRONTGIPSRLEASAS</p>
SEE COUNT	939 a 1139 c 1102 g 1216 t
IGIN	
1	

卷之三

/note="matches EST AA714011 from clone IMAGE:12406921"
complement(2112). .2172)
/note="matches EST BE747194 from clone IMAGE:39295520"
join(277. .254.2548. .2589.2646. .2763)
/note="matches EST AA007343 from clone IMAGE:429263"
complement(2311. .2352)
/note="matches EST AV650272 from clone GLCCDA12
matches EST AV650673 from clone GLCCHC08
matches EST AV650603 from clone GLCGGE10
matches EST AV650460 from clone GLCCEE02
matches EST AV650468 from clone GLCCHD09."
join(2417. .2711.2711. .2814.3001. .3039)
/note="matches EST W029293 from clone IMAGE:291802"
2475. .2804
/note="matches EST W31922 from clone IMAGE:320524"
join(2476. .2815.2813. .2817)
/note="matches EST AA488179 from clone IMAGE:8233660"
join(2647. .2894.2893. .3059.3053. .3115)
/note="matches EST AA4303072 from clone IMAGE:7583437"
join(2739. .3054.3050. .3143)
/note="matches EST AA232187 from clone IMAGE:666656"
join(2757. .2884.2881. .3023)
/note="matches EST N40351 from clone IMAGE:269966"
complement(Join(2794. .2876.2894. .2935.2935. .3045,
3046. .3138.3135. .3249))
/note="matches EST AA007344 from clone IMAGE:429263"
complement(2796. .3249)
/note="matches EST AA133230 from clone IMAGE:1932055"
2808. .2899
/note="matches EST N44116 from clone IMAGE:272953"
complement(Join(2810. .3045.3046. .3250))
/note="matches EST A108631 from clone IMAGE:1660573"
complement(Join(2810. .2885.2885. .3011. .2935.3091. .3250))
/note="matches EST N273535 from clone IMAGE:269966"
complement(Join(2818. .3045.3046. .3092.3091. .3249))
/note="matches EST N36110 from clone IMAGE:272953)
complement(Join(2905. .3045.3046. .3138.3135. .3249))
/note="matches EST AA448918 from clone IMAGE:823360"
complement(Join(2918. .3045.3046. .3138.3135. .3249))
/note="matches EST AA231364 from clone IMAGE:666656"
complement(3040. .3549)
/note="matches EST AW298226 from clone IMAGE:2733069"
join(3212. .3778.3540. .3563)
/note="matches EST W38959 from clone IMAGE:304871"
complement(3397. .3974)
/note="matches EST AA628914 from clone IMAGE:1032940"
3402. .3738
/note="matches EST C04258 from clone 3NHC3019"
join(3459. .3613.3613. .3656)
/note="matches EST AA045035 from clone IMAGE:488780"
3459. .3877
/note="matches EST AA130301 from clone IMAGE:503848"
join(3542. .3710.3774. .3948)
/note="matches EST W39026 from clone IMAGE:305523"
complement(3576. .4126)
/note="matches EST AI042706 from clone IMAGE:1431595"
join(3597. .3809.3798. .4045)
/note="matches EST AA1313497 from clone IMAGE:490637"
complement(Join(3598. .3619.3642. .4038))
/note="matches EST AA115137 from clone IMAGE:490637"
complement(3597. .4126)
/note="matches EST AI041537 from clone IMAGE:1643803"
complement(3608. .3974)
/note="matches EST AI081144 from clone IMAGE:1683131"
complement(3625. .4126)
/note="matches EST AA404352 from clone IMAGE:7583437"
complement(Join(3626. .3645.3674. .3898.3867. .4126))
/note="matches EST AA1313666 from clone IMAGE:503848"
/note="matches EST AW973035"
complement(Join(3685. .3898.3867. .4126))
/note="matches EST AI001151 from clone IMAGE:1680293"
complement(Join(3689. .3898.3867. .4122))

Db	601	TATCCCACTTCAGGAGGTAGGGCCCCAAGCTGGGGGCCACCGGTACTCCCT	660
Qy	902	tctcgacccatccaggcaecgcataaca tgcggccggcggcagtgccctggggc	961
Db	661	TCTGTACCTTCAAGGCAGGATAAACATGCCACCAGTGCCNGGGC	720
Qy	962	tggtgtcttcagaactacaggcagccaaacgtggccaccatct	1021
Db	721	TGGTGTCTTCAGCACTAACAGGGAGCCAAAGTGTGTTGATGCCACCATCT	780
Qy	1022	tca gtc cgg tgg tcc catggggatccatcagccggctgtggggctggg	1081
Db	781	TAGTCCTGGTTGGATTCAGGGATTCAGGGTGCCTGGGTGGCTGGCC	840
Qy	1082	c ag t g a a g g t g c a g t a c e c c t g a c c g t g g a c c g t g a g g	1141
Db	841	CAGTAAGGGTGCAGCTAACCTGACGCCATGGACCTGGTGAAGGCCAGGG	900
Qy	1142	c t c g g t g c a g t c g g c t g t g c c t c a t g c c t g t a g t g c a t a g	1201
Db	901	CTCCTGTCGCTTCATGCCCTCATGCCCTAGGGCATAGGGCTCGTCA	960
Qy	1202	q c t t g c c t g c c a t g g a c t a g g c c a a g t c t c t g t c t g	1261
Db	961	GCTTGTGCCCTGGCCATGGACTCAGGCCAGTGTGCTCATGGCACGGGC	1020
Qy	1262	a g a s a g g c t c c t g g a g a c t t g g c t g t g a g g a c t c c t	1321
Db	1021	AGACGGGCTCCCTGGAGACTCTGGCCCTGTGGAGACTCTCTACCTCCATTCAA	1080
Qy	1322	g i a c c a a t g a g a c c a a a g g a c c a a c t t g c e c a t	1381
Db	1081	GGACCAATAGGACCAATCTGGACACTGCTAGAAACCAAGGCCCATC	1140
Qy	1382	c c a g a t c t g g a g a c c c t c a g g c c o c t c t g g c t c c t	1441
Db	1141	CGATCTGGAAAGCCCCTAGGCCCTCTGGCTGGCCCTCCCTGGGGC	1200
Qy	1442	c c c t c t g c c c g t c g g g g a c t g c a c t g t g c t g t g	1501
Db	1201	CCCTCTCCTCCCTCGGGCATGCACTGGCTGGGCTGACCTGTCGTGATGG	1260
Qy	1502	t c t t g t c a g t c c t c t t c t t c t t c t t c t t c t t c t t	1561
Db	1261	TCTTGTGTCAGTCCTCCCTCCCTGGGCTGGCCAGTGCCTGTCAGCGAGA	1320
Qy	1562	t c t a c c t g t g g a g a t a c g a g a a g a g c c t l c g c t t	1621
Db	1321	TCTACCTGTGAGATACGAGAAAGCTGGCTTCAGCTGGCTGAGTGG	1380
Qy	1622	c c a a c t c t t c t a c g c t c t t c t c g c a t c t t g g e c a	1681
Db	1381	CCAACTCTTCATCACGCTCCTCGATCTGGCTGGGCTCACTAATTTGTC	1440
Qy	1682	c c t t c t g c t c a c g g a c t g c c t c t t c t t c t t c t t	1741
Db	1441	CCCTCTGTCAGTCAGCTGCGCTTCAGCTGGCTGGGCTCACTAATTTGTC	1500
Qy	1742	c t g a a c a a a g g c c a g t c t g g a g a g a t a c e a g c a g	1801
Db	1501	CTGAAACAAAAGGCCAGTCCTGGAGAGATGACCAACTGGCTGGAGACC	1560
Qy	1802	c c t g a g c t t g g c c a c c g g a a g a c t c c a c t g c t	1861
Db	1561	CCCTGAGCTTGGCCACAGGCCAACTCCACTGGCTGGCTCAATTTGTC	1620
Qy	1862	c t g g g c t c t c t g g a a t c c t g c t g g a a c t g y c t t	1920
Db	1621	CTGGGGCTCTCTGAGAATCGCTGGCTGGAGATGACCAACTGGCTGGAGACC	1680
Qy	1921	a t c t c c a g a t c t c t c t a g g c c c a d a g c a a a g t	1980
Db	1681	ATCTCAGCATCAGTCAGCTGGCTGGCTGGAGATGACCAACTGGCTGGAGAST	1740


```

/gene="DJ10IA2.2"
  /note="match: cDNA: Em:AK013049 Em:AK011574
  match: ESTs: Em:AA2337250 Em:T87841 Em:AA113277
    Em:AA98248 Em:AA682750 Em:AA461487 Em:A028588
    Em:AA659710 Em:A1149891 Em:AA579641 Em:AA516077
    Em:AA460132 Em:A1483490 Em:AA016575"
    /evidence="not_experimental"
    /product="dJ10IA2.2 (novel protein)"
    complement("1..3110")
    /gene="dJ10IA2.2"
      18..494
        /note="match: SNS: Em:HS28H20T"
        complement("join(84..562,2463..2745)") )
      CDS
        /gene="dJ10IA2.2"
          /note="supported by FGENESH and GENSCAN
          continues in Em:AL133520 as dJ10IA2.2"
          /codon_start=1
          /evidence="not_experimental"
          /product="dJ28H20.2 (novel protein)"
          /protein_id="CAC00561_1"
          /db_xref="GI:9588402"
          /db_xref="SPTRNBL:QINQE6"
          /translation="MAAA RATT PAGDEPAPAE AELAAA ARERSRFL SGL EKV KOGAE
          ARVERG RGP QGR VAI KHH PFKY RHP ALKEAR LGR RTV DVE RALUR CRAG ISAPV FF
          FVD YAS RNC LKMB IEC OS TVADY IOT MET MTKT PCK QNL AKT GYL A RHM DED LH
          GDL TSNM LKPK PPLE QOL NIVL VDGF GFS FIS ALPE KFL DVY LKE RAFL STH PNT ETIV
          EAFL KSY TS SKK ARPV LKKL DEV RGR KRS MVG"
          repeat_region
            867..1145
              /note="AluSX repeat: matches 1..298 of consensus"
            1251..1353
              /note="MIR repeat: matches 50..151 of consensus"
            1672..1797
              /note="MIR repeat: matches 8..135 of consensus"
            1855..2079
              /note="MIR repeat: matches 33..262 of consensus"
            2155..2248
              /note="MIR repeat: matches 33..135 of consensus"
            2384..3150
              /note="CpG is island"
              /evidence="not_experimental"
            2812..2871
              /note="3 copies 20 mer 85% conserved"
              complement("2876..3110")
            misc_feature
              /note="match: STS: Em:G15621"
              repeat_region
                3965..3962
                  /note="MER21B repeat: matches 299..391 of consensus"
                3963..5655
                  /note="L1A8 repeat: matches 4532..6270 of consensus"
                5669..5895
                  /note="L1NE repeat: matches 5501..5739 of consensus"
                  /note="L332..6332
                    /note="L2 repeat: matches 2524..2746 of consensus"
                    6626..6915
                      /note="AluSX repeat: matches 3..290 of consensus"
                    6918..6980
                      /note="MER94 repeat: matches 1..62 of consensus"
                    7023..7322
                      /note="Alujo repeat: matches 1..312 of consensus"
                    7323..7323
                      /note="L2 repeat: matches 2389..2513 of consensus"
                    7457..7623
                      /note="MIR repeat: matches 93..262 of consensus"
                    7864..7934
                      /note="MLT1 repeat: matches 311..382 of consensus"
                      8046..8240
                        /note="MLT1J repeat: matches 9..211 of consensus"
                      8347..8667
                        /note="AluS repeat: matches 1..300 of consensus"
                      8685..8793
                        /note="MLT1 repeat: matches 271..366 of consensus"
                      9095..9221
                        /note="MIR repeat: matches 102..242 of consensus"
            repeat_region
              9647..9776
                /note="L2 repeat: matches 2170..2303 of consensus"
                9793..10028
                  /note="Alujo repeat: matches 74..306 of consensus"
                10038..10205
                  /note="LUTR16C repeat: matches 126..316 of consensus"
                10631..10690
                  /note="MLT1D repeat: matches 445..505 of consensus"
                10691..10995
                  /note="AluJB repeat: matches 1..295 of consensus"
                10996..11200
                  /note="MLT1D repeat: matches 248..445 of consensus"
                11232..11634
                  /note="MSTB repeat: matches 2..409 of consensus"
                11670..11761
                  /note="MLT1D repeat: matches 128..192 of consensus"
                  /note="MLT1D repeat: matches 1..192 of consensus"
                11762..12062
                  /note="AluSX repeat: matches 1..293 of consensus"
                12063..12180
                  /note="MLT1D repeat: matches 1..128 of consensus"
                  complement("1218L..12344")
                  /note="match: GSS: Em:AO130762"
                13011..13129
                  /note="MLT1J repeat: matches 252..368 of consensus"
                13210..13392
                  /note="MSTB20 repeat: matches 1..215 of consensus"
                13397..13442
                  /note="MIR repeat: matches 104..116 of consensus"
                13660..13783
                  /note="LUTR3 repeat: matches 330..452 of consensus"
                13891..14254
                  /note="L2 repeat: matches 2255..2619 of consensus"
                14305..14435
                  /note="L2 repeat: matches 2617..2748 of consensus"
                14446..14525
                  /note="4 copies 20 mer 76% conserved"
                15476..15752
                  /note="AluB repeat: matches 2..278 of consensus"
                15774..15952
                  /note="MEG58A repeat: matches 42..220 of consensus"
                16190..16304
                  /note="L2 repeat: matches 2657..2744 of consensus"
                16305..16599
                  /note="AluB repeat: matches 1..296 of consensus"
                16600..16635
                  /note="L2 repeat: matches 2617..2657 of consensus"
                16641..16803
                  /note="LIMA10 repeat: matches 6162..6284 of consensus"
                16804..16819
                  /note="MER57B repeat: matches 1..403 of consensus"
                17198..17297
                  /note="5 copies 20 mer 71% conserved"
                17200..17296
                  /note="47 copies 2 mer ta 70% conserved"
                17555..17868
                  /note="MER57B repeat: matches 1..403 of consensus"
                17935..18240
                  /note="LUTR26 repeat: matches 246..553 of consensus"
                  /note="match: SRS: Em:G12885"
                18451..18615
                  /note="FAM repeat: matches 1..312 of consensus"
                19238..19336
                  /note="MLT57-internal repeat: matches 7168..7244 of
                  consensus"
                repeat_region
                  19343..19485
                    /note="MR53 repeat: matches 2..180 of consensus"
                repeat_region
                  19721..20190
                    /note="LUTR8 repeat: matches 251..691 of consensus"
                repeat_region
                  20288..20493
                    /note="LUTR8 repeat: matches 1..256 of consensus"
                repeat_region
                  20704..20978
                    /note="MIR repeat: matches 102..242 of consensus"

```


Center: Joint Genome Institute	JOURNAL	Unpublished	
Center Code: JGI	REFERENCE	2 (bases 1 to 181705)	
Web site: http://www.jgi.doe.gov	AUTHORS	DOE Joint Genome Institute.	
Project Information	TITLE	Direct Submission	
Center Project Name: 600095	JOURNAL	Submitted (18-FEB-2000) Production Sequencing Facility, DOE Joint	
Center clone name: RPCI-11_50412	SEQUENCE	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	
COMMENT	On Apr 25, 2001 this sequence version replaced gi:11178070.	On Apr 25, 2001 this sequence version replaced gi:11178070.	
-----	DEFINITION	-----	
-----	ACCESSION	-----	
-----	VERSION	-----	
-----	KEYWORDS	-----	
-----	ORGANISM	-----	
-----	SOURCE	-----	
-----	FEATURES	-----	
-----	LOCATION	-----	
-----	QUALIFIERS	-----	
-----	BASE COUNT	-----	
-----	ORIGIN	-----	
-----	RESULT	-----	
Summary Statistics	Summary Statistics	Summary Statistics	
Consensus quality: 172472 bases at least Q40	Consensus quality: 169715 bases at least Q40	Consensus quality: 169715 bases at least Q40	
Consensus quality: 174610 bases at least Q30	Consensus quality: 174817 bases at least Q30	Consensus quality: 174817 bases at least Q30	
Estimated insert size: 190110; agarose-fp estimation	Estimated insert size: 175159; sum-of-contigs estimation	Estimated insert size: 176639 bases at least Q20	
Estimated insert size: 10 66 in 020 bases; agarose-fp estimation	Estimated insert size: 217000; agarose-fp estimation	Estimated insert size: 181005; sum-of-contigs estimation	
Quality coverage: 11.57 in Q0 bases; sum-of-contigs estimation	Quality coverage: 9.4 in Q20 bases; agarose-fp estimation	Quality coverage: 9.4 in Q20 bases; agarose-fp estimation	
* NOTE: This is a 'working draft' sequence. It currently	* NOTE: This is a 'working draft' sequence. It currently	* NOTE: This is a 'working draft' sequence. It currently	
* consists of 5 contigs. Gaps between the contigs	* consists of 8 contigs. The true order of the pieces	* consists of 8 contigs. The true order of the pieces	
* are represented as runs of N. The order of the pieces	* is not known and their order in this sequence record is	* is not known and their order in this sequence record is	
* is believed to be correct as given, however the sizes	* arbitrary. Gaps between the contigs are represented as	* arbitrary. Gaps between the contigs are represented as	
* of the gaps between them are based on estimates that have	* runs of N, but the exact sizes of the gaps are unknown.	* runs of N, but the exact sizes of the gaps are unknown.	
* provided by the submittor	* This record will be updated with the finished sequence	* This record will be updated with the finished sequence	
* This sequence will be replaced	* as soon as it is available and the accession number will	* as soon as it is available and the accession number will	
* by the finished sequence as soon as it is available and	* be preserved.	* be preserved.	
* the accession number will be preserved.	* 1 22409: contig of 22409 bp in length	* 1 1212: contig of 1212 bp in length	
* 22410 22509: gap of unknown length	* 22509: gap of unknown length	* 1213 1312: gap of unknown length	
* 31950 31954: contig of 9445 bp in length	* 31954: gap of unknown length	* 1313 2398: contig of 1086 bp in length	
* 31955 32054: gap of unknown length	* 32054: gap of unknown length	* 2399 2498: gap of unknown length	
* 32055 116297: contig of 84243 bp in length	* 116297: contig of 84243 bp in length	* 2499 3678: gap of 1180 bp in length	
* 116298 116397: gap of unknown length	* 116398 145562: contig of 29165 bp in length	* 3678 3679: gap of unknown length	
* 145562 145563: gap of unknown length	* 145563 175599: contig of 29937 bp in length.	* 3679 7829: contig of 4051 bp in length	
* 145563 175599: contig of 29937 bp in length.	* 175599 1 .175599	* 7830 7929: gap of unknown length	
Location/Qualifiers	Source	Location/Qualifiers	
1. .175599	/organism="Homo sapiens"	1. .181705	
/db_xref="Taxon:2606"	/db_xref="Taxon:9606"	/organism="Homo sapiens"	
/chromosome="16"	/clone="RP1-50412"	/db_xref="Taxon:9606"	
/clone="RP1-50412"	/clone lib="RPCI human BAC library 11"	/chromosome="16"	
		/clone lib="RPCI-11 clone BAC library 11"	
		/clone lib="RPCI human BAC library 11"	
BASE COUNT	FEATURES	BASE COUNT	
ORIGIN	source	ORIGIN	
RESULT	source	ORIGIN	
AC023831	Query Match 2.3%; Score 99; DB 2; Length 175599;	AC023831	Query Match 2.3%; Score 99; DB 2; Length 181705;
LOCUS	Best Local Similarity 100.0%; Pred. No. 1.7e-4;	LOCUS	Best Local Similarity 100.0%; Pred. No. 1.7e-4;
DEFINITION	Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	DEFINITION	Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ACCESSION	SEQUENCE, 8 unorderd pieces.	ACCESSION	SEQUENCE, 8 unorderd pieces.
VERSION	AC023831.7	VERSION	AC023831.7
KEYWORDS	HGVS; HTGS; DRAFT; HTGS_ACTIVEFIN.	KEYWORDS	HGVS; HTGS; DRAFT; HTGS_ACTIVEFIN.
ORGANISM	Homo sapiens	ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 181705)	REFERENCE	1 (bases 1 to 181705)
AUTHORS	DOI Joint Genome Institute.	AUTHORS	DOI Joint Genome Institute.
TITLE	Sequencing of Human Chromosome 16	TITLE	Sequencing of Human Chromosome 16

LOCUS	AC023105	70876 bp	DNA	PRI	09-MAY-2001
DEFINITION	Homo sapiens BAC clone GS1-13M19 from 7, complete sequence.				
VERSION	AC023105				
KEYWORDS	HTG.				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS	Muiyane, E., Nguyen, C., Strownatt, C. and Waligorski, J.				
TITLE	JOURNAL				
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1. (bases 1 to 70876)				
AUTHORS	Sulston, J. E. and Waterston, R.				
TITLE	JOURNAL				
JOURNAL	Toward a complete human genome sequence				
REFERENCE	2. (bases 1 to 70876)				
AUTHORS	Muiyane, E., Nguyen, C., Strownatt, C. and Waligorski, J.				
TITLE	JOURNAL				
JOURNAL	The sequence of Homo sapiens BAC clone GS1-13M19 Unpublished 3. (bases 1 to 70876)				
REFERENCE	Waterston, R. H.				
AUTHORS	Direct Submission				
TITLE	JOURNAL				
JOURNAL	Submitted (08-FEB-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
REFERENCE	4. (bases 1 to 70876)				
AUTHORS	Waterston, R. H.				
TITLE	JOURNAL				
JOURNAL	Submitted (29-NOV-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
REFERENCE	5. (bases 1 to 70876)				
AUTHORS	Waterston, R.				
TITLE	JOURNAL				
JOURNAL	Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
COMMENT	On Nov 29, 2000 this sequence version replaced g1:9838027.				
-----	Genome Center				
-----	Center: Washington University Genome Sequencing Center				
-----	Center code: WUGSC				
-----	Web site: http://genome.wustl.edu/gsc				
-----	Contact: apliens@wustl.edu				
-----	Summary Statistics				
-----	Center project name: H_GS013M19				
-----	-----				
NOTICE:	This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.				
	This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.				
MAPPING INFORMATION:	The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see http://www.ncbi.nlm.nih.gov/DIR/GB/CHR/ , send mailto: green@nigri.nih.gov , or see http://genome.wustl.edu/gsc				
SOURCE INFORMATION:	This clone is from the first BAC library from Genome Systems, Inc. (http://www.genomesystems.com). Call us: 1-800-Lymphoblastoid two				

FEATURES
source
 misc_feature
 51108: contig of 4504 bp in length
 51109: gap of unknown length
 51208: contig of 4386 bp in length
 51209: gap of unknown length
 55594: contig of 6068 bp in length
 55595: gap of unknown length
 61762: contig of 6068 bp in length
 61763: gap of unknown length
 61863: contig of 8109 bp in length
 69971: gap of unknown length
 70072: contig of 9183 bp in length
 79254: gap of unknown length
 79354: contig of 14856 bp in length
 94210: gap of unknown length
 94211: contig of 16952 bp in length
 94311: gap of unknown length
 111262: contig of 21504 bp in length
 111363: gap of unknown length
 132867: gap of unknown length
 132966: gap of unknown length
 132967: contig of 26783 bp in length
 159750: gap of unknown length
 159850: contig of 26077 bp in length.

Location/Qualifiers
 1..185926
 Organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-487E1"
 note="assembly_name:Contig16"
 1..1472
 note="assembly_name:Contig17"
 1573..2852
 note="assembly_name:Contig18"
 2953..4741
 note="assembly_name:Contig19"
 4842..8040
 note="assembly_name:Contig19"
 8141..110050
 note="assembly_name:Contig20"
 11151..11348
 note="assembly_name:Contig21"
 14449..16009
 note="assembly_name:Contig22"
 clone_end:77
 vector_size:left"
 18110..22446
 note="assembly_name:Contig23"
 22447..22763
 note="assembly_name:Contig24"
 27364..31414
 note="assembly_name:Contig25"
 31151..36552
 note="assembly_name:Contig26"
 36353..41629
 note="assembly_name:Contig27"
 41730..46504
 note="assembly_name:Contig28"
 46605..51108
 note="assembly_name:Contig29"
 51209..55594
 note="assembly_name:Contig30"
 55695..61762
 note="assembly_name:Contig31"
 61863..69971
 note="assembly_name:Contig32"
 70072..79254
 note="assembly_name:Contig33"
 79355..94210
 note="assembly_name:Contig34"
 94311..111262
 note="assembly_name:Contig35"
 111363..132866
 note="assembly_name:Contig36"
 clone_end:SP6
 vector_size:right"
 132967..159749
 note="assembly_name:Contig37"

 misc_feature
 519850..185926
 note="assembly_name:Contig38"
 BASE COUNT 57022 a 36205 c 36048 g 54445 t 2206 others
 ORIGIN
 Query Match 1..58; Score 64; DB 2; Length 185926;
 Best Local Similarity 100.0%; Pred. No. 8 1e-24;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 64;
 Query 2380 ctgatctggctactcaacctccacttctgggtcaaggatcttcgtcgcc 2439
 Db 82937 CTGATCTGGCAGACTGCAACCTCCACTTCCNGGTCAGGGATTCCTGCCTCGGCC 82996
 Query 2440 tcct 2443
 Db 82997 TCCT 83000
 RESULT 13
 AC020950
 LOCUS AC020960 108162 bp DNA HTG 16-JUL-2000
 DEFINITION Mus musculus clone CT7-240L13, WORKING DRAFT SEQUENCE, 1 ordered
 pieces.
 ACCESSION AC020960.2 GI:92563183
 KEYWORDS HTGS_PHASE; HTGS_DRAFT.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scluognathi; Murinae; Mus .
 REFERENCE 1 (bases 1 to 108162)
 AUTHORS DOE Joint Genome Institute.
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 108162)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
 COMMENT Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Jul 18, 2000 this sequence version replaced gi:6691270.

 Project Information
 Center Project Name: 1045618
 Center clone name: RG-MBAC-240L13

 Summary Statistics
 Consensus quality: 103750 bases at least Q40
 Consensus quality: 106923 bases at least Q30
 Consensus quality: 107282 bases at least Q20
 Estimated insert size: 108212; sum-of-contigs estimation
 Quality coverage: 6.53 In Q20 bases; pulse field gel estimation
 Quality coverage: 7.85 In Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced.
 * The finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 108162: content of 108162 bp in length.

 FEATURES
 Source
 1..108162
 Organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="CT7-240L13".

BASE COUNT 34815 /clone_1.lib="CitrinC7 mouse BAC library"
ORIGIN a 21949 c 21217 g 30178 t 3 others

* is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 * 1 3553: contig of 3553 bp in length
 * 3654 3655: gap of unknown length
 * 3654 8755: contig of 5102 bp in length
 * 8756 8755: gap of unknown length
 * 8856 35675: contig of 26820 bp in length
 * 35676 35775: gap of unknown length
 * 85484: contig of 49709 bp in length
 * 85585: gap of unknown length
 * 85585 94427: contig of 8843 bp in length
 * 94428 94522: gap of unknown length
 * 94528 105891: contig of 11364 bp in length
 * 105892 105991: gap of unknown length
 * 10592 112271: contig of 6280 bp in length
 * 112272 112371: gap of unknown length
 * 112372 113982: contig of 1611 bp in length.
 FEATURES
 source
 1..113982
 /organism="Homo sapiens"
 /db_xref="TAXON:9606."
 /chromosome="8"
 /map="8q11-12."
 /clone="CTA-204B4."
 BASE COUNT 28505 a 28727 c 27506 g 28544 t 700 others
 ORIGIN
 Query Match 1..4% Score 63; DB 2; Length 113982;
 Best Local Similarity 100.0%; Pred. No. 2.8e-23;
 Matches 63; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;
 Qy 2381 tgatctggtcaccaactccatctctgggttcaacgcattttccgtccatcgct 2440
 Db 101068 TGATCTGGTCACTCAGCATTCCACTGCAAGCTCCATGGTTCAAGCGATCTGCCNAGCC 101127
 Oy 2441 cct 2443
 Db 101128 CCT 101130
 RESULT 14
 AF192303 113982 bp DNA HTG 07-JUN-2001
 LOCUS Homo sapiens chromosome 8 clone CTA-204B4 map 8q11-12, WORKING
 DEFINITION DRAFT SEQUENCE, 8 unordered pieces.
 ACCESSION AF192303
 VERSION GI:14327853
 KEYWORDS HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Human.
 ORGANISM Homo sapiens
 Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo.
 Schilhabel, M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
 Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudt,A.,
 Siddiqui,R., Taudien,S., Wen,G., Korenberg,J., Rosenthal,A. and
 Platzer,M.
 Chromosome 8 genomic sequence
 Unpublished 1 (bases 1 to 113982)
 Schilhabel, M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
 Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudt,A.,
 Siddiqui,R., Taudien,S., Wen,G., Korenberg,J., Rosenthal,A.
 Platzer,M.
 Direct Submission
 Submitted (04-OCT-1999) Genome Analysis, Institute of Molecular
 Biotechnology, Beutenbergstrasse 11, Jena 0745, Germany
 On Jun 7, 2001 This sequence version replaced gi:8151740.
 Center: Institute of Molecular Biotechnology
 Center code: IMB
 Web site: http://genome.imb-jena.de/
 Contact: gscj-submit@genome.imb-jena.de
 ----- Project Information
 Center project name: H211
 Center clone name: CTA-204B4
 ----- Summary Statistics
 Sequencing vector: M13; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.99029
 Consensus quality: 106890 bases at least Q40
 Consensus quality: 109881 bases at least Q30
 Consensus quality: 111616 bases at least Q20
 Quality coverage: 5.74 x in 020 bases; sum-of-contigs
 Sequence Quality Assessment:
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality 10.
 Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.
 * NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces

Center Project Name: 704688
 Center clone name: CITB-H1_220003

.....

Summary Statistics

Consensus quality: 116661 bases at least Q40
 Consensus quality: 120243 bases at least Q30
 Consensus quality: 121387 bases at least Q20
 Estimated insert size: 125000; pulse field gel estimation
 Estimated insert size: 121972; sum-of-contigs estimation
 Quality coverage: 8.51 in Q20 bases; pulse field gel estimation
 Quality coverage: 8.72 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 9 contigs.
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.

This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

- * 36662: contig of 36662 bp in length
- * 36762: gap of unknown length
- * 36763: 45137: contig of 8375 bp in length
- * 45138: 45327: gap of unknown length
- * 45238: 54176: contig of 8939 bp in length
- * 54177: 54277: gap of unknown length
- * 54277: 85219: contig of 31443 bp in length
- * 85720: 85819: gap of unknown length
- * 85820: 88481: contig of 2662 bp in length
- * 88482: 88581: gap of unknown length
- * 88582: 93330: contig of 4749 bp in length
- * 93331: 93330: gap of unknown length
- * 93431: 96395: contig of 3065 bp in length
- * 96496: 96596: gap of unknown length
- * 96596: 119670: contig of 23075 bp in length
- * 119671: 119770: gap of unknown length
- * 119771: 122772: contig of 3002 bp in length.

FEATURES
 source
 1. 122772

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="5"

/clone="CRD-20003"

/clone.lib="Caltech human BAC library D"
 25066 c 23097 g 34312 t 811 others

BASE COUNT
 ORIGIN

Query Match Score 63; DB 2; Length 122772;
 Best Local Similarity 100.0%; Pred. No. 2.8e-23;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2381 tgatctggctactgcaacctccacttctcgtggttcacgatttcgtccatgcgtcagcct 2440
 Db 120510 TGATCTGGCTACTGCAACCTCCACTTGCTCAAGCGATTCTCTGGTTCAAGCGATTCTCTGGCTCAGCCT 120569
 Oy 2441 oct 2443
 Db 120570 CTC 120572

Search completed: March 15, 2002, 07:03:55
 Job time: 10176 sec

THIS PAGE BLANK (USPTO)